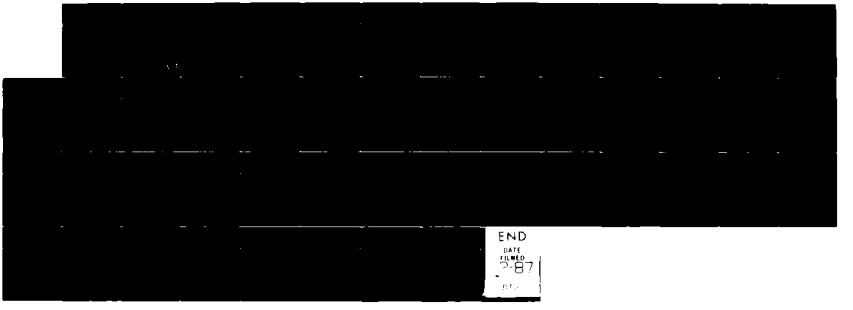
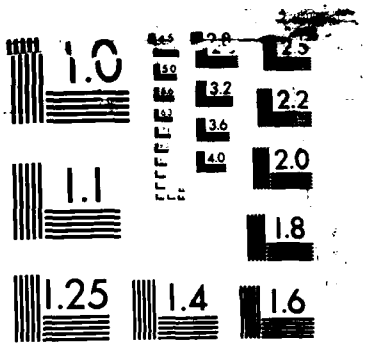
AD-A170 004 ANALYSIS OF FAULT TOLERANT COMPUTER SYSTEMS(U) ARIZONA 1/1
UNIV TUCSON DEPT OF MATHEMATICS U SUMITA ET AL. FEB 85
AFOSR-TR-86-0403 AFOSR-84-0205

UNCLASSIFIED F/G 9/2 NL



END
DATE FILMED
7-87



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS

AD-A170 004

SECURITY CLASSIFICATION

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION: Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY: N/A	3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for Public Release; Distribution Unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE: N/A	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S): AFOSR-TR- 86-0403
6a. NAME OF PERFORMING ORGANIZATION: Department of Mathematics, University of Arizona; 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION: AFOSR/NM
6c. ADDRESS (City, State and ZIP Code): Tucson, Arizona 85721	7b. ADDRESS (City, State and ZIP Code): Bldg. 410, Bolling AFB, DC 20332-6448
8a. NAME OF FUNDING/SPONSORING ORGANIZATION: U.S. Air Force (AFOSR); 8b. OFFICE SYMBOL (If applicable): NM	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: AFOSR - 84 - 0205
8c. ADDRESS (City, State and ZIP Code): AFOSR/NM, Building 410, Bolling AFB, DC 20332 - 6448	10. SOURCE OF FUNDING NOS.

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.
6.1102F	2304	A5	

11. TITLE (Include Security Classification): Analysis of Fault Tolerant Computer Systems

12. PERSONAL AUTHOR(S): Ushio Sumita, J. George Shanthikumar and Yasushi Masuda

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
Technical report	FROM _____ TO _____	February 1985	45

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB. GR.	Fault tolerant computer systems, availability, reliability, performability, computational procedures.

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

In many critical applications of digital systems, fault tolerance has been an essential architectural attribute for achieving high reliability. In recent years, the concept of the performability of such systems has drawn the attention of many researchers. In this paper, we develop a general Markov model for fault tolerant computer systems. Various important performance measures, including the performability measures as well as some new performance measures, are treated in a unified manner. Furthermore general and efficient computational procedures are developed for calculating these performance measures based on the uniformization technique of Keilson (1974, 1979). A numerical example is given to illustrate the computational procedures developed.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐	Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE NUMBER (Include Area Code)	22c. OFFICE SYMBOL
Major Brian W. Woodruff	(202) 767-5027	AFOSR/NM

Analysis of Fault Tolerant Computer Systems

by

U. Sumita
J. G. Shanthikumar†
Y. Masuda

Working Paper Series No. QM8507

February 1985

The Graduate School of Management
The University of Rochester

†School of Business Administration
University of California at Berkeley,
Berkeley, California 94720

The first author was partially supported by the National Science Foundation under Grant No. ECS-8404071. The second author was partially supported by Air Force office of Scientific Research, U.S.A.F., under Grant AFOSR-84-0205. Reproduction in whole or in part is permitted for any purpose of the United States Government.

ABSTRACT

In many critical applications of digital systems, fault tolerance has been an essential architectual attribute for achieving high reliability. In recent years, the concept of the performability of such systems has drawn the attention of many researchers. In this paper, we develop a general Markov model for fault tolerant computer systems. Various important performance measures, including the performability measures as well as some new performance measures, are treated in a unified manner. Futhermore general and efficient computational procedures are developed for calculating these performance measures based on the uniformization technique of Keilson(1974,1979). A numerical example is given to illustrate the computational procedures developed.

§0 Introduction

In many critical applications of digital systems such as flight control, nuclear plant control, etc., the need for achieving high reliability has made fault-tolerance an essential architectual attribute of digital systems. In general, to achieve high reliability requirements, some redundancy techiniques are employed where systems contain multiple copies of a resource. Typically forms of redundant structures for fault-tolerant systems are categorized into four classes and combinations thereof, see Beaudry(1978). In Massive Redundant Systems, redandunt techniques such as triple-modular redundancy (Von Neumann(1956)), N-modular redundancy(Mathur and Avizienis(1970)), and self-purging redundancy(Losq(1976)) are employed where the same task is executed on each equivalent module and the vote on the outputs is taken for improving the output information. In Standby Redundant Systems, tasks are executed on active units in the system. When a failure of an active unit is detected, the system attempts to replace the faulty unit with a spare unit, see Bouricius, Coorter, Jessep and Schneider(1969). Hybrid Redundant Systems consist of massive redundant cores with spares to replace failed modules, see Losq(1976). In Gracefully Degrading Systems, all operative units in the system are kept active for executing tasks. Upon the detection of a unit failure the system attempts to reconfigure the remaining operative units and continue operation, see Borgesson and Freitas(1975). The reader is referred to an exellent paper by Avizienis(1978) for a more thorough discussion on the concept of fault-tolerance in digital systems and a chronological view of the evolution of fault-tolerant systems.

A substantial literature exists for developing and analyzing reliability models of these fault-tolerant digital systems, see e.g. Arnold (1973), Beaudry (1978), Borgerson and

Freitas (1975), Bouricius, Coorter, Jessp and Schneider (1969), Castillo and Siewiorek (1981), Costes, Landrault and Laprie (1978), Gay and Ketelson (1979), Huslende (1981), Iyer, Donatiello and Heidelberger (1984), Koren and Sue (1979), Krishna and Shin (1983), Makam and Avizienis (1981,1982), Mathur and Avizienis (1970), Meyer (1980,1982), Meyer and Furchtgott and Wu (1980), Ng and Avizienis (1977,1980), Oda, Tohma and Furuya (1981), Osaki and Nishino (1980), Seth and Lipsky (1983), Sonerio and Suk (1980), Trivedi (1982) and others. Validation of models for such high reliable systems has also been discussed, see e.g. Trivedi, Gault and Clery(1980). No general computational schemes, however, have been developed for evaluating important reliability measures.

The purpose of this paper is three-fold. First, a general stochastic model for fault-tolerant computer systems is developed. Underlying distributions of interest can be system state dependent, incorporating possible interdependency among multiple modules. The distributions are not restricted to exponential distributions. The model is general in that any Markov chain model in the literature can be viewed as a special case, including the unified reliability model of Ng and Avizenis(1980), and provides substantial modeling flexibility for the performance analysis of such systems. Second, various important performance measures (some of them are new) are treated in a unified manner. In particular, several performability measures are discussed concerning the computational capacity of the system in the time interval $[0, t)$. Finally efficient and general computational procedures are developed for evaluating all of these performance measures, using the uniformization technique of Keilson(1974,1979).

To the auther's best knowedge, the concept of the performability of the system can be traced back to early 60's. The distribution and the moments of functionals of Markov

renewal processes was studied by Jewell(1963) and Mclean and Neuts(1967). Relevant weak and strong laws were also examined by Pyke and Schanfale(1964). In an excellent and substantial paper by Çinlar(1969), functionals of semi-Markov processes were examined where the transform of the time dependent distributions of such functionals and recursion formulae for the moments thereof were established. Keilson and Rao(1970,1971) studied the limiting behavior as $t \to \infty$ of processes defined on Markov chain, having state dependent growth rate.

Recently the concept of the performability of the system has been revitalized in the context of fault tolerant computer systems, see Meyer(1980,1982). Iyer, Donatiello and Heidelberger(1984) developed a recursion formulae in a Markov chain context using the spectral representation. These recent papers failed to provide the reference to the relevant previous work described above. The results of Çinlar(1969) were derived based on renewal type arguments. In this paper, we provide an independent and totally analytic proof for the Markov chain case. An extension of this analytic proof to semi-Markov case is straightforward.

In Section 1, a general stochastic model will be developed for fault tolerant computer systems. The model enables one to incorporate time dependent availability, reliability and performability measures of such systems. These performance measures will be classified, in Section 2, into three categories: (A) availability and reliability measures independent of computational capacity; (B) performability measures involving cumulative computational capacity; (C) performability measures involving computational capacity during the first passage time to the system failure. Most of performance measures in the category (A) are traditional. The performance measures in the category (B) are con-

cerned with the time dependent performability of the system which have drawn attention of researchers during the last few years. One of performance measures in the category (C) was first introduced and analyzed by Beaudry(1978). We will extend this work in a more systematic manner. Section 3 through 5 will be devoted to develop general computational schemes for evaluating the performance measures in the categories (A) through (C) respectively. Finally in Section 6, a numerical example will be given, illustrating the efficiency of the computational procedures developed.

§.1 Model Description

We consider a fault tolerant computer system consisting of N modules. The module i may be in any one of $K_i + 1$ states $S_i = \{0, 1, \ldots, K_i\}$. Some of these states in S_i may represent the operative state of the module and the remaining represent the failed states of the module under repair. Let $X_i(t)$ be the state of the module i at time t. We also define the indicator process $I(t)$ by

$$I(t) = \begin{cases} 0, & \text{if the system is not functioning at time } t; \\ 1, & \text{if the system is functioning at time } t. \end{cases} \tag{1.1}$$

Then the multivariate process $\underline{J}(t)$ defined by

$$\underline{J}(t) = [X_1(t),\, X_2(t), \ldots, X_N(t),\, I(t)] \tag{1.2}$$

fully describes the state of the system at time t. The state space of $\underline{J}(t)$ is given by

$$S = S_1 \times S_2 \times \ldots \times S_N \times \{0, 1\}. \tag{1.3}$$

We note that the indicator process $I(t)$ is deliberately introduced. One may have a system where, given a state vector $\underline{x} = (x_1, \ldots, x_N)$ of N modules, the system is operative only with certain probability.

We assume that the multivariate process $\underline{J}(t)$ is a temporally homogeneous Markov process on S governed by the transition rate matrix

$$\underline{\underline{\nu}} = [\nu_{\underline{m}\,\underline{n}}], \quad \underline{m}, \underline{n} \in S. \tag{1.4}$$

It should be noted that the underlying distributions are not restricted to exponential distributions. Futhermore N modules are not necessarily mutually independent. For example, suppose that the up-time of the module i is an Erlang-2 random variable. Upon a failure, it takes a random duration exponentially distributed before the repair

starts. The repair time itself is also exponentially distributed. One then sets S_i = {0, 1, 2, 3} where 0 represents the failed state under repair; 1 represents the failed state with no repair; 2 represents the operative state under the second phase of the up-time; and 3 represents the operative state under the first phase of the up-time. Moreover parameter values involved could depend on the states of other modules. Hence the model could incorporate Phase-type distributions of Neuts(1981) with possible interdependence among modules. Thus the model considered here is quite general and provides substantial modeling flexibility for the performance analysis of fault tolerant computer systems.

The computational capacity of the whole system can be characterized by a mapping $\phi : S \rightarrow R^+$. The value $\phi(\underline{J}(t))$ represents the maximum amount of computation per unit time that the system can provide at time t. We decompose the state space S into two subsets G and B where

$$G = \{\underline{m} : \underline{m} \in S,\ m_{N+1} = 1\} \quad \text{and} \quad B = S \setminus G \,. \tag{1.5}$$

The subset G is called a good set since $\phi(\underline{m}) > 0$ for any $\underline{m} \in G$. Similarly the subset B is called a bad set since $\phi(\underline{m}) = 0$ for any $\underline{m} \in B$.

In some applications, it may be of interest to study the computational capacity of the system for specific jobs. We assume that a set of jobs, C, to be processed by the system consists of M different classes $C_j, 1 \leq j \leq M$. As before the computational capacity of the system for the jobs in C_j is characterized by a mapping $\phi_j : S \rightarrow R^+$. It should be noted, however, that modules required for processing jobs may vary depending on classes. Hence it is possible to have $\phi_j(\underline{m}) = 0$ for some $\underline{m} \in G$. Accordingly we also define a good set G_j and a bad set B_j for each class C_j, i.e.,

$$G_j = \{\underline{m} : \underline{m} \in S,\ \phi_j(\underline{m}) > 0\} \quad \text{and} \quad B_j = S \setminus G_j \,. \tag{1.6}$$

We note that $G_j \subset G$ and $B \subset B_j$ for all j.

The decomposition of the state space into subsets G and B together with computational capacity function ϕ enables one to describe the system behavior more accurately. In particular, it is often assumed in the modeling of multiprocessor digital systems that the failed state of the system due to the failure in coverage and the failed state of the system due to the failure of all modules are the same, see e.g. the state 0 in Figure 7 of Beaudry(1978). In our model, this distinction can be made clearly, allowing one to introduce different distributions for recovery times. This point will be illustrated through a numerical example in Section 6.

The transition rate matrix $\underline{\underline{\nu}}$ of (1.4) depends heavily on the system structure. By specifying $\underline{\underline{\nu}}$, any Markov chain model appeared in the literature can be viewed as a special case of this model. In the next section, we introduce various performance measures concerning availability, reliability and performability of the system for a general transition rate matrix $\underline{\underline{\nu}}$ and general computational capacity function ϕ and ϕ_j. The general computational schemes for calculating these performance measures will be also developed in the following sections.

§.2 Performance Measures

Several different traditional and performance related reliability measures have been considered for fault tolerant computer systems in the literature. These performance measures are often used to compare alternative configuration for fault tolerance. In this section, we present both time dependent and stationary performance measures concerning availability, reliability and performability of the general model developed in Section 1. This provides a concise summary of key performance measures discussed in the literature. Furthermore some important new performance measures are also introduced. The performance measures we consider in this section are classified into three categories.

(A) Availability and reliability measures independent of computational capacity

These measures are intended to provide information about the state of the system and have no relevance to computational capacity of the system at different system states.

(A1) State probability

Let $\underline{a}^T$ be the initial state probability vector for the multivariate process $\underline{J}(t)$, i.e.,

$$\underline{a}^T = (a_{\underline{m}})_{\underline{m}\in S}\,; \qquad a_{\underline{m}} = P[\underline{J}(0) = \underline{m}], \quad \underline{m} \in S. \tag{2.1}$$

The state probability vector at time t given $\underline{a}^T$ is clearly of interest. We denote this vector by

$$\underline{p}^T(t|\underline{a}) = \left(p_{\underline{m}}(t|\underline{a})\right)_{\underline{m}\in S}\,; \qquad p_{\underline{m}}(t|\underline{a}) = P[\underline{J}(t|\underline{a}) = \underline{m}], \quad \underline{m} \in S, \tag{2.2}$$

where $\underline{J}(t|\underline{a})$ denotes the state of the system at time t given the initial state probability vector $\underline{a}^T$. When there is no confusion, we will write $\underline{p}^T(t)$ instead of $\underline{p}^T(t|\underline{a})$. When the transition rate matrix $\underline{\underline{\nu}}$ is irreducible, the ergodic state probability vector exists where

$$\underline{e}^T = \lim_{t\to\infty} \underline{p}^T(t|\underline{a}). \tag{2.3}$$

(A2) Point and interval availability

The point availability $A(t|\underline{a})$ at time t is the probability that the system is operational at time t given $\underline{a}^T$. The interval availability $AI(t,\tau|\underline{a})$ at time t is the expected fraction of interval $(t,t+\tau)$ during which the system is operational. Formally we define

$$A(t|\underline{a}) = P[\underline{J}(t|\underline{a}) \in G] \tag{2.4}$$

and

$$AI(t,\tau|\underline{a}) = \frac{1}{\tau} E\left[\int_t^{t+\tau} I(x|\underline{a})dx\right] \tag{2.5}$$

where $I(x|\underline{a})$ is the indicator function of (1.1) given $\underline{a}$. At ergodicity, one has

$$A_\infty = \lim_{t\to\infty} A(t|\underline{a}) = \lim_{t\to\infty} AI(t,\tau|\underline{a}) = AI_\infty\,. \tag{2.6}$$

(A3) Time to first system failure and related reliability

Suppose that $\underline{J}(0) = \underline{\theta}$ where $\underline{\theta} \in G$. Of interest is the time until the first system failure. This is the first passage time of the multivariate process $\underline{J}(t)$ from $\underline{\theta} \in G$ to the bad states B, defined by

$$T_{\underline{\theta}\,B} = \inf\{t : \underline{J}(t) \in B \mid \underline{J}(0) = \underline{\theta} \in G\}. \tag{2.7}$$

If $\underline{J}(t)$ has an initial state probability vector $\underline{a}^T$, then the corresponding first passenge time $T_{\underline{a}\,B}$ would be a probability mixture of $T_{\underline{\theta}\,B}$ weighted by $a_{\underline{\theta}}$ where $T_{\underline{\theta}\,B} = 0$ with probability one for $\underline{\theta} \in B$. Typically the system is operative at time $t = 0$ and one has $a_{\underline{\theta}} = 0$ for $\underline{\theta} \in B$. Otherwise $T_{\underline{a}\,B}$ has mass $\sum_{\underline{\theta}\in B} a_{\underline{\theta}}$ at the origin.

Of interest is the cumulative distribution function of $T_{\underline{a}\,B}$ defined by

$$F_{\underline{a}\,B}(x) = P[T_{\underline{a}\,B} \le x]. \tag{2.8}$$

An important reliability measure is the probability that the system will continue to be operative for the period longer than x given $\underline{a}^T$. We denote this reliability measure by $R_{\underline{a}}(x)$. One then has

$$R_{\underline{a}}(x) = P[T_{\underline{a}\,B} > x] = 1 - F_{\underline{a}\,B}(x). \tag{2.9}$$

Also of interest are the moments

$$E[T^k_{\underline{a}\,B}], \quad k = 1, 2, \ldots, \tag{2.10}$$

and the α-reliable mission time τ_α defined by

$$\tau_\alpha = \sup\{x : R_{\underline{a}}(x) \geq \alpha R_{\underline{a}}(0+)\}. \tag{2.11}$$

We note that the first moment $E[T_{\underline{a}\,B}]$ is the mean time to failure.

(A4) Time to next system failure at time t and interval reliability

Given an initial state probability vector $\underline{a}^T$, let $T_{\underline{a}\,B|t}$ be the time to the next system failure from time t if the system is up at time t and zero otherwise. We define

$$F_{\underline{a}\,B|t}(x) = P[T_{\underline{a}\,B|t} \leq x]. \tag{2.12}$$

We note that $T_{\underline{a}\,B|t}$ has mass $F_{\underline{a}\,B|t}(0+) = 1 - A(t|\underline{a})$ at the origin. Of interest is an interval reliability $RI_{\underline{a}|t}(x)$ given by

$$RI_{\underline{a}|t}(x) = P[T_{\underline{a}\,B|t} > x] = 1 - F_{\underline{a}\,B|t}(x). \tag{2.13}$$

$RI_{\underline{a}|t}(x)$ is the probability that, given $\underline{a}$, the system will continue to operate until time $t + x$ from time t. Corresponding to (2.10) and (2.11), we are interested in the moments

$$E[T^k_{\underline{a}\,B|t}], \quad k = 1, 2, \ldots, \tag{2.14}$$

and the α-reliable interval mission time

$$\tau_{\alpha|t} = sup\{x : RI_{\underline{a}|t}(x) \geq \alpha RI_{\underline{a}|t}(0+)\}. \tag{2.15}$$

In some applications, one may be interested in the conditional random variable $T_{\underline{a}\,B|t}$ given that $T_{\underline{a}\,B|t} > 0$. We denote this conditional random variable by $T^{+}_{\underline{a}\,B|t}$. Reliability measures in (2.12) through (2.15) can be modified accordingly in a straightforward manner. For example, one has, corresponding to (2.13),

$$RI^{+}_{\underline{a}|t}(x) = P[T^{+}_{\underline{a}\,B|t} > x] = RI_{\underline{a}|t}(x)/A(t|\underline{a}). \tag{2.16}$$

Here $RI^{+}_{\underline{a}|t}(x)$ is the probability that the system will continue to be operative until time $t + x$ given that it is operative at time t, with the initial probability vector $\underline{a}$.

(A5) Stationary reliability measures

When the transition rate matrix $\underline{\underline{\nu}}$ is irreducible, the multivariate Markov process $\underline{J}(t)$ is ergodic. Hence the random variable $T_{\underline{a}\,B|t}$ converges in distribution to a random variable, say S, as $t \to \infty$. The limiting random variable S denotes the time until next failure at ergodicity. We see that

$$F_S(x) = P[S \leq x] = \lim_{t\to\infty} F_{\underline{a}\,B|t}(x). \tag{2.17}$$

With probability $F_S(0+) = 1 - A_\infty$, the system is not functioning at ergodicity. Hence S has mass $1 - A_\infty$, at the origin. Stationary reliability measures corresponding to (2.13) through (2.15) can be found by letting $t \to \infty$. In particular we define the stationary interval reliability $RI_S(x)$ by

$$RI_S(x) = P[S > x] = \lim_{t\to\infty} RI_{\underline{a}|t}(x). \tag{2.18}$$

and the α-reliable stationary mission time τ^S_α by

$$\tau^S_\alpha = \sup\{x : RI_S(x) \geq \alpha RI_S(0+)\}. \tag{2.19}$$

The conditional random variable $T^{+}_{\underline{a}\,B|t}$ also converges in distribution to $S^{+} \stackrel{\mathrm{d}}{=} S_{|S>0}$ as $t \to \infty$. Stationary reliability measures for S^{+} can be obtained similarly. One has, for example,

$$RI^{+}_{S}(x) = P[S^{+} > x] = RI_{S}(x)/A_{\infty}. \tag{2.20}$$

(A6) Quasi-stationary reliability measures

Suppose that the system has been operating for a "long time". One then wishes to know how long it will take from current time until the first system failure. As for the stationary case, the conditional random variable of $T_{\underline{a}\,B|t}$ given that $\underline{J}(0) \in G$ and $T_{\underline{a}\,B} > t$ converges in distribution to a random variable, say Q, as $t \to \infty$. This limiting random variable Q is called the quasi-stationary exit time from G, see e.g. Keilson(1974, 1979). More formally we define

$$F_{Q}(x) = P[Q \leq x] = \lim_{t\to\infty} P[T_{\underline{a}\,B|t} \leq x | T_{\underline{a}\,B} > t,\, \underline{J}(0) \in G]. \tag{2.21}$$

Quasi-stationary reliability measure can then be introduced in terms of $F_{Q}(x)$. In particular we define the quasi-stationary interval reliability $RI_{Q}(x)$ by

$$RI_{Q}(x) = P[Q > x] = 1 - F_{Q}(x), \tag{2.22}$$

and the α-reliable quasi-stationary mission time τ^{Q}_{α} by

$$\tau^{Q}_{\alpha} = \sup\{x : RI_{Q}(x) \geq \alpha\}. \tag{2.23}$$

(A7) Cumulative operational time during the interval $[0,t)$

Another important reliability measure of interest is the cumulative operational time of the system during the time period $[0,t)$ given the initial state probability vector $\underline{a}^{T}$.

We denote this random variable by $CO(t|\underline{a})$. One then has

$$CO(t|\underline{a}) = \int_0^t I(x|\underline{a})dx. \tag{2.24}$$

Evaluation of the distribution of $CO(t|\underline{a})$ is quite hard. We will derive the expression of the moments

$$E[CO(t|\underline{a})^k], \quad k = 1, 2. \tag{2.25}$$

Associated computational procedure will be also developed. It should be noted that $E[CO(t|\underline{a})] = tAI(0, t|\underline{a})$.

One may also consider several other compound reliability measures such as the joint measure for the number of system failures during $[0, t)$ and the system state at time t, see e.g. Baxter(1982), Masuda, Shanthikumar and Sumita(1984), Shanthikumar(1983) and Sumita and Shanthikumar(1984). It should be noted that the reliability measures described in (A5) and (A6) have not been discussed in the context of fault tolerant computer systems.

(B) Performance Measures Involving Cumulative Computational Capacity

When the computational capacity of the system in operational state is constant independent of the actual state of the system (as in the case of identical standby redundant systems), all the performance measures described in Section A2 can be directly related to computational capacity measures. However this is not the case in every system. As described in Section 0, the gracefully degrading system reacts to a detected failure by reconfigurating the system modules, which leads to a new system state possibly with a decreased level of performance. The performance measures to be discussed in this section are concerned with the computational capacity of the system in a finite time interval.

(B1) Cumulative computational capacity in the time interval $[0,t)$

Let $V(t|\underline{a})$ be the cumulative computational capacity of the system for the whole class C of jobs in the interval $[0,t)$ given $\underline{a}$. More formally we define

$$V(t|\underline{a}) = \int_0^t \phi\left(\underline{J}(x|\underline{a})\right) dx \tag{2.26}$$

where $\phi : S \rightarrow R^+$ is the computational capacity function introduced in Section 1 and $\underline{J}(t|\underline{a})$ denotes the state of the system at time t given $\underline{a}$. As mentioned earlier, the modules required for processing may vary depending on classes of jobs. Hence it may be desirable to study the computational capacity of the system for C_j jobs. We denote this random variable by $V_j(t|\underline{a})$ where

$$V_j(t|\underline{a}) = \int_0^t \phi_j\left(\underline{J}(x|\underline{a})\right) dx. \tag{2.27}$$

(B2) Cumulative computational capacity in the time interval $[t,t+\tau)$

Of related interest is the cumulative computational capacity of the system in the time interval $[t,t+\tau)$, $\tau > 0$. Following the notation of (B1) we define

$$V(t,\tau|\underline{a}) = \int_t^{t+\tau} \phi\left(\underline{J}(x|\underline{a})\right) dx \tag{2.28}$$

and

$$V_j(t,\tau|\underline{a}) = \int_t^{t+\tau} \phi_j\left(\underline{J}(x|\underline{a})\right) dx. \tag{2.29}$$

Çinlar(1969) established the transform $E[e^{-wV(t|\underline{a})}]$ explicitly and provided a recursion formula for calculating the moments of $V(t|\underline{a})$ in the semi-Markov context. The computational scheme for the moments of $V(t|\underline{a})$ has been developed in a recent paper by Iyer, Donatiello and Heidelberger(1984) using the spectal representation of the underlying Markov chain. In section 4, we will provide an independent derivation of the double

transform $\int_0^\infty e^{-st}E[e^{-wV(t|\underline{a})}]dt$. Numerical procedures for calculating the first two moments of the performability measures in (2.26) through (2.29) will be also developed.

(C) Performance Measures Involving Computational Capacity during the First Passage Time to System Failure

Performance-related reliability measures involving computational capacity of a computer system during the first passage time to system failure were first studied in Beaudry (1978). In what follows, we describe these together with ~~special~~ other related reliability measures. We discuss only the total computational capacity of the system. The computational capacity of jobs in the class C_j can be studied in a similar manner, where the mapping $\phi : S \to R^+$ should be repalced by $\phi_j : S \to R^+$. For the future reference, we indicate this by adding the index j to the expressions for the total computational capacity.

(C1) Computational capacity before the first system failure and computational reliability

Given $\underline{a}$, let $W_{|\underline{a}}$ be the computational capacity available from the system before the first system failure. More formally, one has

$$W_{|\underline{a}} = \int_0^{T_{\underline{a}B}} \phi(\underline{J}(x|\underline{a}))dx. \tag{2.30}$$

We denote the distribution function of $W_{|\underline{a}}$ by

$$F_{W|\underline{a}}(x) = P[W_{|\underline{a}} \leq x]. \tag{2.31}$$

Suppose that a task requiring x units of computational time is initiated at time $t = 0$. Then the probability that this task will be computed without any interruptions due to system failure is given by $P[W_{|\underline{a}} > x] = 1 - F_{W|\underline{a}}(x)$. We call this measure a computational

reliability denoted by $R_{W|\underline{a}}(x)$ i.e.

$$R_{W|\underline{a}}(x) = 1 - F_{W|\underline{a}}(x). \tag{2.32}$$

Of related interest are the moments

$$E[W^k_{|\underline{a}}], \quad k = 1, 2, \ldots \tag{2.33}$$

and the α-reliable task length defined by

$$\ell_\alpha = \sup\{x : R_{W|\underline{a}}(x) \geq \alpha\}. \tag{2.34}$$

We note that ℓ_α is the maximum computational length of a task that has a probability of α or more for being completed before the first system failure.

(C2) Time dependent computational capacity until next system failure and interval computational reliability

Let $W_{|\underline{a},t}$ be the total computational capacity available from the system from time t until the next system failure if the system is operative at time t, and zero otherwise. That is,

$$W_{|\underline{a},t} = \int_0^{T_{\underline{p}(t|\underline{a})B}} \phi\left(\underline{J}(x|\underline{p}(t|\underline{a}))\right) dx. \tag{2.36}$$

The corresponding cumulative distribution function is denoted by

$$F_{W|\underline{a},t}(x) = P[W_{|\underline{a},t} \leq x]. \tag{2.37}$$

The interval computational reliability $RI_{W|\underline{a}t}(x)$ is then defined by

$$RI_{W|\underline{a},t}(x) = P[W_{|\underline{a},t} > x] = 1 - F_{W|\underline{a},t}(x). \tag{2.38}$$

$RI_{W|\underline{a},t}(x)$ is the probability that the system is operative at time t and it will successfully complete a task of computational length x before its next failure. The counterparts of

(2.33) and (2.34) for $W_{\underline{a},t}$ are

$$E[W^k_{|\underline{a},t}], \quad k = 1, 2, \ldots \tag{2.39}$$

and

$$\ell_{\alpha|t} = \sup\{x : RI_{W|\underline{a},t}(x) \geq \alpha RI_{W|\underline{a},t}(0+)\} \tag{2.40}$$

Here $\ell_{\alpha|t}$ is the α-reliable interval task length, representing the maximum computational length of a task that will be completed before the next system failure with probability α or more, if it is initiated at time t. We note that $W_{\underline{a}|t}$ has mass $P[\underline{J}(t|\underline{a}) \in B]$ at the origin. The corresponding measure associated with the conditional random variable $W^+_{\underline{a}|t} = W_{\underline{a}|t|W_{\underline{a}|t}>0}$ can be studied following the argument in (A4).

(C3) Stationary computational measures

When the system is ergodic, the random variable $W_{\underline{a}|t}$ converges in distribution to a random variable, say S_W, as $t \to \infty$. The limiting random variable S_W denotes the computational capacity of the system until the next system failure at ergodicity. One sees that

$$F_{S_W}(x) = P[S_W \leq x] = \lim_{t\to\infty} F_{W|\underline{a},t}(x). \tag{2.41}$$

We note that S_W has mass $1 - A_\infty$ at the origin, i.e. $F_{S_W}(0+) = 1 - A_\infty$. Stationary computational measures corresponding to (2.38) to (2.40) can be found by letting $t \to \infty$. We define

$$RI_{S_W}(x) = P[S_W > x] = 1 - F_{S_W}(x) \tag{2.42}$$

and

$$\ell^S_\alpha = \sup\{x : RI_{S_W}(x) \geq \alpha RI_{S_W}(0+)\}. \tag{2.43}$$

The conditional measures associated with $S^+ = S_W|_{S_W>0}$ can be discussed similary.

(C4) Quasi-stationary computational measures

The quasi-stationary reliability measures associated with the quasi-stationary exit time Q was discussed in (A6). In this section we examine the quasi-stationary computational measures. Suppose that the system has been operating for a long time. The question to be answered is how large the computational capacity of the system would be before the next system failure. Formally the random variable Q_W denoting the above quantity can be defined by the limiting distribution of $W_{\underline{a}|t}$ given that $\underline{J}(0) \in G$ and $T_{\underline{a}\,B} > t$ as $t \to \infty$. That is,

$$F_{Q_W}(x) = P[Q_W \leq x] = \lim_{t\to\infty} P[W_{\underline{a}|t} \leq x | T_{\underline{a}\,B} > t,\ \underline{J}(0) \in G] \tag{2.44}$$

Quasi-stationary computational measure can then be introduced through $F_{Q_W}(x)$. In particular, we define the quasi-stationary interval computational reliability $RI_{Q_W}(x)$ by

$$RI_{Q_W}(x) = P[Q_W > x] = 1 - F_{Q_W}(x) \tag{2.45}$$

and the α-reliable quasi-stationary task length ℓ_α^Q by

$$\ell_\alpha^Q = \sup\{x : RI_{Q_W}(x) \geq \alpha\}. \tag{2.46}$$

The computational capacity measures described in (C3) and (C4) have not been discussed in the literature.

§3. Numerical Procedures for Computing Performance Measures in Category (A)

As we will see, all performance measures described in Section 2(A) can be expressed in terms of the time dependent state probability vector $\underline{p}^T(t)$ of the underlying Markov chain $\underline{J}(t)$, possibly with certain modifications. Hence for the computation of these performance measures, it is necessary to develop efficient numerical procedures to evaluate $\underline{p}^T(t|\underline{a})$. In the next subsection, we show that the uniformization procedures of Keilson(1974, 1979) provides the computational vechicle needed for this purpose.

3.1 State probability

We have assumed that the underlying process $\underline{J}(t)$ is a finite Markov chain in continuous time on S governed by transition rate matrix $\underline{\underline{\nu}} = [\nu_{\underline{m}\,\underline{n}}]$. Let $\underline{\underline{p}}(t) = [p_{\underline{m}\,\underline{n}}(t)]$ be the transition probability matrix of $\underline{J}(t)$, that is

$$p_{\underline{m}\,\underline{n}}(t) = P[\underline{J}(t) = \underline{n} | \underline{J}(0) = \underline{m}], \quad \underline{m}, \underline{n} \in S. \tag{3.1}$$

Let

$$\nu_{\underline{m}} = \sum_{\underline{n} \in S, \underline{n} \neq \underline{m}} \nu_{\underline{m}\,\underline{n}}. \tag{3.2}$$

Since the cardinality of S denoted by $|S|$ is finite, there exists a positive ν such that

$$\sup_{\underline{m} \in S} \nu_{\underline{m}} \leq \nu. \tag{3.3}$$

A Markov chain in continuous time is said to be uniformizable if its governing transition rates satisfy (3.3). All finite Markov chains in continuous time are automatically uniformizable. Keilson(1974, 1979) has shown that the uniformizability provides a useful bridge between continuous time Markov chains and discrete time Markov chains in the following manner.

For notational convenience, we define the diagonal matrix $\underline{\underline{\nu}}_D$ whose diagonal elements are $\nu_{\underline{m}}$'s ordered appropriately. From the Kolmogorov forward equations, one then has

$$\underline{\underline{p}}(t) = e^{\underline{\underline{Q}}t}; \quad \underline{\underline{Q}} = -\underline{\underline{\nu}}_D + \underline{\underline{\nu}}. \tag{3.4}$$

Here $\underline{\underline{Q}}$ is the infinitesimal generator. Using the uniformization constant ν defined in (3.3), we define

$$\underline{\underline{a}}_\nu = \underline{\underline{I}} - \frac{1}{\nu}\underline{\underline{\nu}}_D + \frac{1}{\nu}\underline{\underline{\nu}} \tag{3.5}$$

where $\underline{\underline{I}}$ is an identity matrix of size $|S|$. We note that $\underline{\underline{a}}_\nu \geq 0$ and $\underline{\underline{a}}_\nu \underline{1} = \underline{1}$ where $\underline{1}$ is the vector of length $|S|$ having all elements equal to 1. Hence $\underline{\underline{a}}_\nu$ is a stochastic matrix. From (3.4) one sees that $\underline{\underline{Q}} = -\nu[\underline{\underline{I}} - \underline{\underline{a}}_\nu]$ so that

$$\underline{\underline{p}}(t) = e^{-\nu t[\underline{\underline{I}} - \underline{\underline{a}}_\nu]} = \sum_{k=0}^{\infty} q_k(t)\underline{\underline{a}}_\nu^k \tag{3.6}$$

where

$$q_k(t) = e^{-\nu t}\frac{(\nu t)^k}{k!}, \quad k = 0, 1, 2, \ldots. \tag{3.7}$$

Here $\underline{\underline{a}}_\nu^0 = \underline{\underline{I}}$. Equation (3.6) provides a bridge between a Markov chain in discrete time governed by $\underline{\underline{a}}_\nu$ and the Markov chain in continuous time governed by $\underline{\underline{\nu}}$. Hence given an initial state probability vector $\underline{a}$, one has

$$\underline{p}^T(t|\underline{a}) = \sum_{k=0}^{\infty} q_k(t)\underline{a}^T\underline{\underline{a}}_\nu^k. \tag{3.8}$$

Equation (3.8) enables one to calculate $\underline{p}^T(t|\underline{a})$ efficiently via computer. Although the expression involves an infinite series, the matrix norm of $\underline{\underline{a}}_\nu^k$ is bounded by one for all k and the truncation point k^* may be determined for a given accuracy ε and sufficiently large $T > 0$ by

$$k^* = \min\{k : \sum_{j=0}^{k} q_j(t) > 1 - \varepsilon, \quad 0 \leq t \leq T\}. \tag{3.9}$$

The ergodic probability vector $\underline{e}^T$ can be found by letting $t \to \infty$ in (3.8). Alternatively $\underline{e}^T$ may be found by solving $\underline{e}^T \underline{\underline{a}}_\nu = \underline{e}^T$ and $\underline{e}^T \underline{1} = 1$.

3.2 Point and interval availability

The point availability $A(t|\underline{a})$ defined in (2.4) can be expressed straightforwardly in terms of $\underline{p}^T(t|\underline{a})$. We denote the subvector of $\underline{b}$ of length $|S|$ restricted to the good set G by $\underline{b}_G$. One then sees that

$$A(t|\underline{a}) = \underline{p}_G^T(t|\underline{a})\underline{1}_G. \tag{3.10}$$

For the interval availability $AI(t,\tau) = \frac{1}{\tau}E[\int_t^{t+\tau} I(x|\underline{a})dx]$ in (2.5), one first observes from the linearity of $E[\cdot]$ and the boundedness of integral that $AI(t,\tau) = \frac{1}{\tau}\int_t^{t+\tau} E[I(x|\underline{a})]dx$. Since $E[I(x|\underline{a})] = P[\underline{J}(x|\underline{a}) \in G]$, one has

$$AI(t,\tau|\underline{a}) = \frac{1}{\tau}\int_t^{t+\tau} \underline{p}_G^T(x|\underline{a})\underline{1}_G dx. \tag{3.11}$$

The integral $\int_t^{t+\tau} \underline{p}_G^T(x|\underline{a})dx$ can be computed from (3.8) as

$$\int_t^{t+\tau} \underline{p}^T(x|\underline{a})dx = \sum_{k=0}^{\infty} Q_k(t,\tau)\underline{a}^T \underline{\underline{a}}_\nu^k \tag{3.12}$$

where

$$Q_{k+1}(t,\tau) = \frac{1}{\nu}(q_{k+1}(t) - q_{k+1}(t+\tau)) + Q_k(t,\tau), \quad k = 0,1,2\ldots, \tag{3.13}$$

starting with $Q_0(t,\tau) = \frac{1}{\nu}e^{-\nu t}(1 - e^{-\nu\tau})$.

3.3 Time to first system failure and related reliability

To find the distribution of the first passage time $T_{\underline{a}\,B}$, we consider the absorbing process $\underline{J}^*(t)$ obtained from the original process $\underline{J}(t)$ by censoring transitions from B to G, see Keilson(1979). It is easily seen that the infinitesimal generator $\underline{\underline{Q}}^*_{G\,G}$ governing $\underline{J}^*(t)$ inside the set G is given by

$$\underline{\underline{Q}}^*_{G\,G} = -\underline{\underline{\nu}}_{D:G\,G} + \underline{\underline{\nu}}_{G\,G} \tag{3.14}$$

where $\underline{\underline{b}}_{G\,G}$ denotes the submatrix of a $|S| \times |S|$ matrix restricted to G. Correspondingly, the transition probability matrix $\underline{\underline{p}}^{*}_{G\,G}(x)$ of $\underline{J}^{*}(t)$ is given by

$$\underline{\underline{p}}^{*}_{G\,G}(x) = exp\{\underline{\underline{Q}}^{*}_{G\,G}x\} = \sum_{k=0}^{\infty} q_k(x)\underline{\underline{a}}^{k}_{\nu:G\,G}. \tag{3.15}$$

Since the reliability $R_{\underline{a}}(x)$ of (2.9) is given by $R_{\underline{a}}(x) = P[T_{\underline{a}\,B} > x] = P[\underline{J}^{*}(x|\underline{a}) \in G]$, one has

$$R_{\underline{a}}(x) = \sum_{k=0}^{\infty} q_k(x)\underline{a}^{T}_{G}\underline{\underline{a}}^{k}_{\nu:G\,G}\underline{1}_{G} \tag{3.16}$$

and

$$F_{\underline{a}\,B}(x) = P[T_{\underline{a}\,B} \leq x] = 1 - R_{\underline{a}}(x). \tag{3.17}$$

We note that if $\underline{a}^{T}_{G}\underline{1}_{G} < 1$, then $T_{\underline{a}\,B}$ has mass $(1 - \underline{a}^{T}_{G}\underline{1}_{G})$ at the origin. For the moments of $T_{\underline{a}\,B}$, one has (see Keilson (1974, 1979) or Neuts(1981))

$$E[T^{k}_{\underline{a}\,B}] = \frac{k!}{\nu^{k}}\underline{a}^{T}_{G}\underline{\underline{z}}^{k}_{G\,G}\underline{1}_{G} \tag{3.18}$$

where $\underline{\underline{z}}_{G\,G}$ is the fundamental matrix defined by

$$\underline{\underline{z}}_{G\,G} = [\underline{\underline{I}}_{G\,G} - \underline{\underline{a}}_{\nu:G\,G}]^{-1}. \tag{3.19}$$

The α-reliable mission time τ_{α} of (2.11) can be found from (3.16).

3.4 Time to next system failure at time t and interval reliability

For given $\underline{a}^{T}$, the random variable $T_{\underline{a}\,B|t}$ was defined in (A4) as the time until next system failure from time t if the system is operative at time t, and zero otherwise. Because of the Markov property of the underlying process $\underline{J}(t)$, this random variable is equal in distribution to the ordinary first passage time having the initial state probability vector $\underline{p}^{T}(t|\underline{a})$, i.e.

$$T_{\underline{a}\,B|t} \stackrel{\mathrm{d}}{=} T_{\underline{p}(t|\underline{a})\,B}. \tag{3.20}$$

Hence performance measures (2.12) through (2.15) involving $T_{\underline{a}\,B|t}$ can be readily obtained from (3.16) through (3.18) where $\underline{a}_G^T$ should be replaced by $\underline{p}_G^T(t|\underline{a})$. As noted in 2.(A4) $T_{\underline{a}\,B|t}$ has mass $1 - A(t|\underline{a})$ at the origin. The conditional performance measures involving $T^+_{\underline{a}\,B|t} \stackrel{\mathrm{d}}{=} T_{\underline{a}\,B|t}$ given $T_{\underline{a}\,B|t} > 0$ can be found accordingly.

3.5 Stationary reliability measures

We have seen that the random variable $T_{\underline{a}\,B|t}$ converges in distribution to S as $t \to \infty$ when the process $\underline{J}(t)$ is ergodic. It can be seen from (3.20) that $S \stackrel{\mathrm{d}}{=} T_{\underline{e}\,B}$. Hence the stationary reliability measures described in 2.(A5) can be computed using (3.16) through (3.18) where $\underline{a}^T$ should be replaced by $\underline{e}^T$. The conditional measures can be calculated accordingly.

3.6 Quasi-stationary reliability measures

For the quasi-stationary exit time Q introduced in 2.(A6), we assumed that the good set G is irreducible, i.e. all states in G can communicate each other within G. Under this condition, it is known that Q is exponentially distributed, see Keilson(1974, 1979). More specifically one has

$$F_Q(x) = 1 - e^{-\nu(1-\lambda_Q)x} \tag{3.21}$$

where λ_Q is the maximum eigenvalue of the matrix $\underline{a}_{\nu:G\,G}$. λ_Q may be found either by solving a set of equations or by the power method. Then the quasi-stationary reliability measures can be computed straightforwardly.

The cumulative operational time $CO(t|\underline{a})$ described in 2.(A7) can be viewed as a special case of cumulative computational capacity of the system, $V(t|\underline{a})$ discussed in 2.(B1), where $\phi(\underline{m}) = 1$, $\underline{m} \in G$. Hence numerical procedures for finding the moments

of $CO(t|\underline{a})$ can be found from those for finding the moments of $V(t|\underline{a})$, which we discuss next.

§.4. Numerical Procedures for Computing Performance Measures in Category (B)

The results of Çinlar(1969) derived based on renewal type arguments enable us to develop numerical procedures for evaluating computational capacity of the system described in Section 2(B). In this section, however, we derive the results through an independent analytic approach. An extension of the proof to semi-Markov case is straightforward and is omitted here. The moments of the performability measures are then placed in a form for which the uniformization technique can be readily applied.

4.1 Cumulative computational capacity in the time interval $[0,t)$

In order to evaluate cumulative computational capacity of the system in the interval $[0,t)$, we consider the process $Z(t)$ defined by

$$\frac{d}{dt}Z(t)|_{\underline{J}(t)=\underline{m}} = \gamma_{\underline{m}}, \quad \underline{m} \in S \tag{4.1}$$

where $\gamma_{\underline{m}} \geq 0$. The process $Z(t)$ increases at the rate of $\gamma_{\underline{m}}$ while the underlying process $\underline{J}(t)$ is in state $\underline{m}$. We note that if the initial state probability vector of $\underline{J}(t)$ is $\underline{a}$, then one has

$$Z(t) = CO(t|\underline{a}) \quad \text{if } \gamma_{\underline{m}} = 1, \quad \underline{m} \in G, \text{ and } \gamma_{\underline{m}} = 0, \text{ otherwise} \tag{4.2}$$

$$Z(t) = V(t|\underline{a}) \quad \text{if } \gamma_{\underline{m}} = \phi(\underline{m}), \quad \underline{m} \in S \tag{4.3}$$

$$Z(t) = V_j(t|\underline{a}) \quad \text{if } \gamma_{\underline{m}} = \phi_j(\underline{m}), \quad \underline{m} \in S. \tag{4.4}$$

Let

$$F_{\underline{m}}(x,t) = P[Z(t) \leq x, \underline{J}(t) = \underline{m}], \quad \underline{m} \in S. \tag{4.5}$$

We assume that the initial distributions are given by

$$F_{\underline{m}}(x,0) = a_{\underline{m}} D_{\underline{m}}(x) \tag{4.6}$$

where $D_{\underline{m}}$ are an absolutely continuous cumulative distribution functions having probability density functions $d_{\underline{m}}(x)$. It can be easily seen that $F_{\underline{m}}(x,t)$ are then also absolutely continuous and one can define

$$f_{\underline{m}}(x,t) = \frac{\partial}{\partial x} F_{\underline{m}}(x,t), \quad \underline{m} \in S. \tag{4.7}$$

We note that $f_{\underline{m}}(x,t) = 0$ if $x < 0$ or $t < 0$. The p.d.f. of $Z(t)$ denoted by $f(x,t)$ is then given by

$$f(x,t) = \sum_{\underline{m} \in S} f_{\underline{m}}(x,t). \tag{4.8}$$

For the event $\{\underline{J}(t) = \underline{m},\ Z(t) = x\}$ to occur, either $\underline{J}(t)$ starts at $\underline{m}$ and remains there or $\underline{J}(t)$ starts somewhere, enters $\underline{m}$ at time y, $0 < y < t$, and stays in $\underline{m}$ for the period (y,t). By examining the probabilistic flow of the bivariate process $(\underline{J}(t), Z(t))$ in this way, one finds that

$$f_{\underline{m}}(x,t) = f_{\underline{m}}(x - \gamma_{\underline{m}} t, 0) e^{-\nu_{\underline{m}} t} + \sum_{\underline{i} \in S} \nu_{\underline{i}\,\underline{m}} \int_0^\infty f_{\underline{i}}(x - \gamma_{\underline{m}} y, t - y) e^{-\nu_{\underline{m}} y} dy, \quad \underline{m} \in S. \tag{4.9}$$

For convenience, we apply the uniformization technique of Keilson(1974, 1979) here. Using the uniformization constant ν and the associated stochastic matrix $\underline{a}_\nu$ of (3.5), Equation (4.9) can be written as

$$f_{\underline{m}}(x,t) = f_{\underline{m}}(x - \gamma_{\underline{m}} t, 0) e^{-\nu t} + \nu \sum_{\underline{i} \in S} a_{\nu:\underline{i}\,\underline{m}} \int_0^\infty f_{\underline{i}}(x - \gamma_{\underline{m}} y, t - y) e^{-\nu y} dy, \quad \underline{m} \in S. \tag{4.10}$$

Let $\hat{\varphi}_{\underline{m}}(w,t) = \int_0^\infty e^{-wx} f_{\underline{m}}(x,t) dx$ and $\hat{\hat{\varphi}}_{\underline{m}}(w,s) = \int_0^\infty \int_0^\infty e^{-wx-st} f_{\underline{m}}(x,t) dx\, dt$. By taking the double transform of (4.10) one obtains

$$\hat{\hat{\varphi}}_{\underline{m}}(w,s) = \frac{1}{s + \nu + \gamma_{\underline{m}} w} \left[\hat{\varphi}_{\underline{m}}(w,0) + \nu \sum_{\underline{i} \in S} a_{\nu:\underline{i}\,\underline{m}} \hat{\hat{\varphi}}_{\underline{i}}(w,s) \right], \quad \underline{m} \in S. \tag{4.11}$$

Equation (4.11) can be expressed more succinctly in matrix notation. We define the transform vector by $\underline{\hat{\varphi}}(w,s) = [\hat{\varphi}_{\underline{m}}^T(w,s)]_{\underline{m}\in S}$. The vector $\underline{\hat{\varphi}}^T(w,0)$ is defined similarly. Let $\underline{\underline{\gamma}}_D$ be the diagonal matrix having diagonal elements $\gamma_{\underline{m}}$ ordered appropriately and define $\underline{\underline{V}}(w,s) = (s+\nu)\underline{\underline{I}} + w\underline{\underline{\gamma}}_D$. One then sees from (4.11) that

$$\underline{\hat{\varphi}}^T(w,s)\underline{\underline{V}}(w,s) = \underline{\varphi}^T(w,0) + \nu\underline{\hat{\varphi}}^T(w,s)\underline{\underline{a}}_\nu. \tag{4.12}$$

It can be easily seen that $(\underline{\underline{V}}(w,s) - \nu\underline{\underline{a}}_\nu)^{-1}$ exists for $Re(w) > 0$, $Re(s) > 0$. Furthermore, one has $\underline{\underline{\pi}}(s) = \int_0^\infty e^{-st}\underline{\underline{p}}(t)dt = [(s+\nu)\underline{\underline{I}} - \nu\underline{\underline{a}}_\nu]^{-1} = [\underline{\underline{V}}(w,s) - \nu\underline{\underline{a}}_\nu - w\underline{\underline{\gamma}}_D]^{-1}$. Hence from (4.12) we obtain

$$\underline{\hat{\varphi}}^T(w,s) = \underline{\varphi}^T(w,0)[\underline{\underline{I}} + w\pi(s)\underline{\underline{\gamma}}_D]^{-1}\underline{\underline{\pi}}(s). \tag{4.13}$$

Let $\hat{\chi}(w,s) = \int_0^\infty e^{-st}E[e^{-wZ(t)}]dt = \int_0^\infty\int_0^\infty e^{-wx-st}f(x,t)dx\,dt$. Then $\hat{\chi}(w,s) = \underline{\hat{\varphi}}^T(w,s)\underline{1}$ from (4.8). Since $\underline{\underline{\pi}}(s)\underline{1} = \underline{1}/s$, this then leads to

$$\hat{\chi}^T(w,s) = \frac{1}{s}\underline{\varphi}^T(w,s)[\underline{\underline{I}} + w\underline{\underline{\pi}}(s)\underline{\underline{\gamma}}_D]^{-1}\underline{1}. \tag{4.14}$$

If $Z(0+) = 0$, we choose a sequence of $(D_{nj}(x))_{j=0}^\infty$ so that $D_{nj}(x) \to U(x)$ as $j \to \infty$ where $U(x) = 1$, $x \geq 0$ and $U(x) = 0$, $x < 0$. Correspondingly $\underline{\varphi}_j^T(w,0) \to \underline{a}^T$ as $j \to \infty$. One then has

$$\hat{\chi}^T(w,s) = \frac{1}{s}[1 + \sum_{k=1}^\infty (-1)^k w^k \underline{a}^T\{\underline{\underline{\pi}}(s)\underline{\underline{\gamma}}_D\}^k\underline{1}]. \tag{4.15}$$

Althrough the double inversion of (4.15) is quite awkward, it does provide the moment formula. One easily sees that

$$\int_0^\infty e^{-st}E[Z^k(t)]dt = \frac{k!}{s}\underline{a}^T(\underline{\underline{\pi}}(s)\underline{\underline{\gamma}}_D)^k\underline{1}, \quad k = 1,2,\ldots \tag{4.16}$$

or equivalently

$$E[Z^k(t)] = k!\int_0^t \underline{a}^T\{\underline{\underline{p}}(y)\underline{\underline{\gamma}}_D\}^{(k)}\underline{1}dy, \quad k = 1,2,\ldots \tag{4.17}$$

where $\{\underline{\underline{p}}(t)\underline{\underline{\gamma}}_D\}^{(k+1)} = \int_0^t \underline{\underline{p}}(t-y)\underline{\underline{\gamma}}_D\{\underline{\underline{p}}(y)\underline{\underline{\gamma}}_D\}^{(k)}dy$. Using the uniformization procedure described in 3.1, we have:

$$E[Z(t)] = \sum_{n=0}^{\infty} Q_n(0,t)\underline{a}^T\underline{\underline{a}}_{\nu}^n\underline{\underline{\gamma}}_D\underline{1} \tag{4.18}$$

$$E[Z^2(t)] = 2\sum_{m=0}^{\infty}\sum_{n=0}^{\infty} Q_{m+n}(0,t)\underline{a}^T\underline{\underline{a}}_{\nu}^m\underline{\underline{\gamma}}_D\underline{\underline{a}}_{\nu}^n\underline{\underline{\gamma}}_D\underline{1}. \tag{4.19}$$

We note that $\lim_{t\to\infty} E[Z^k(t)]/t^k = \underline{e}^T\underline{\underline{\gamma}}_D^k\underline{1}$. Keilson and Rao(1970,1971) have shown that $\{Z(t) - E[Z(t)]\}/\sqrt{\text{Var } Z(t)}$ converges in distribution to $N(0,1)$ as $t \to \infty$.

§.5 Numerical Procedures for Computing Performance Measures in Category (C)

Given a task requiring certain computational time, it is of interest to find the probability that the task can be processed without interruption due to system failure. In order to answer this question, the distribution of the cumulative computational capacity before the first machine failure is needed. This random variable can be denoted by $Z(T_{\underline{a}B})$ where $Z(t)$ is defined in (4.1). In this section, we develop numerical procedure for calculating performance measures involving $Z(T_{\underline{a}B})$ by employing the trick used in Beaudry(1978) in a more systematic manner.

5.1 Computational capacity before the first system failure and computational reliability

The distribution and the moments of $Z(T_{\underline{a}B})$ can be obtained from the results in 3.3 by modifying the transition rate matrix $\underline{\nu}$ in the following manner. If the process $\underline{J}(t)$ enters state $\underline{m}$, it stays there for an exponentially distributed period with parameter $\nu_{\underline{m}}$. Upon the expiration of the period, the process changes its state to $\underline{n}$ with probability $\nu_{\underline{m}\,\underline{n}}/\nu_{\underline{m}}$, $\underline{n} \in S$. During this dwell time in state $\underline{m}$, the process $Z(t)$ increases by an exponentially distributed amount with parameter $\nu_{\underline{m}}/\gamma_{\underline{m}}$, where this increment is understood to be zero if $\gamma_{\underline{m}} = 0$. Suppose we consider an alternative process $\tilde{\underline{J}}(t)$ on S such that when it enters state $\underline{m}$, it stays there for an exponentially distributed time with parameter $\tilde{\nu}_{\underline{m}} = \nu_{\underline{m}}\,\gamma_{\underline{m}}$. As before, this dwell time is zero if $\gamma_{\underline{m}} = 0$. At the termination of this period, it moves to the state $\underline{n}$ with probability $\nu_{\underline{m}\,\underline{n}}/\nu_{\underline{m}}$. It is then clear that the sequence of the states visited by $\tilde{\underline{J}}(t)$ and $\underline{J}(t)$ are probabilistically the same provided that the two processes share the same initial probability vector $\underline{a}$. However the time required for $\tilde{\underline{J}}(t)$ to achieve

such transitions gives the accumulated computational capacity during those transitions.

If $\gamma_{\underline{m}} > 0$ for any $\underline{m} \in G$ (c.f. cases in (4.2) and (4.3)), one sees that $\underline{\tilde{J}}(t)$ is a Markov chain on S governed by

$$\underline{\underline{\tilde{\nu}}} = \begin{pmatrix} \underline{\underline{\gamma}}_{D:GG}^{-1}\underline{\underline{\nu}}_{GG} & \underline{\underline{\gamma}}_{D:GG}^{-1}\underline{\underline{\nu}}_{GB} \\ \underline{\underline{\nu}}_{BG} & \underline{\underline{\nu}}_{BB} \end{pmatrix}. \tag{5.1}$$

The diagonal matrix $\underline{\underline{\tilde{\nu}}}_D$ is defined accordingly. It can then be readily seen that

$$\tilde{T}_{\underline{a}B} \stackrel{\mathrm{d}}{=} Z(T_{\underline{a}B}) \tag{5.2}$$

where

$$\tilde{T}_{\underline{a}B} = \inf\{t : \underline{\tilde{J}}(t) \in B | \underline{J}(0) = \underline{m} \text{ with probability } a_{\underline{m}},\ \underline{m} \in S\}. \tag{5.3}$$

Hence the distribution and the moments of $Z(T_{\underline{a}B})$ can be obtained from (3.14) through (3.19) where $\underline{\underline{\nu}}_{D:GG}$ and $\underline{\underline{\nu}}_{GG}$ should be replaced by $\underline{\underline{\tilde{\nu}}}_{D:GG}$ and $\underline{\underline{\tilde{\nu}}}_{GG}$. Accordingly

$$\underline{\underline{\tilde{a}}}_{\nu:GG} = \underline{\underline{I}}_{GG} - \frac{1}{\nu}\underline{\underline{\tilde{\nu}}}_{D:GG} + \frac{1}{\nu}\underline{\underline{\tilde{\nu}}}_{GG} \tag{5.4}$$

and

$$P[\tilde{T}_{\underline{a}B} > x] = \sum_{k=0}^{\infty} q_k(x)\underline{a}_G^T\underline{\underline{\tilde{a}}}_{\nu:GG}^k\underline{1}_G. \tag{5.5}$$

For the stationary computational measures described in Section 2(C3), the initial distribution $\underline{a}^T$ in (5.5) should be replaced by the ergodic vector $\underline{e}^T$. For the quasi-stationary computational measures of Section 2(C4), one has to calculate the quasi-stationary vector $\underline{q}_G^T$ on G first. The vector $\underline{q}_G^T$ can be found, for example, using the matrix $\underline{\underline{a}}_{\nu:GG}$ and the power method since $\underline{q}_G^T\underline{\underline{a}}_{\nu:GG} = \lambda_Q\underline{q}_Q^T$. The vector $\underline{a}_G^T$ in (5.5) should be replaced by $\underline{q}_Q^T$.

When $\gamma_{\underline{m}} = 0$ for some $\underline{m} \in G$ (such as the case (4.4)), the state $\underline{m}$ becomes an instantaneous state for $\tilde{\underline{J}}(t)$. That is the process $\tilde{\underline{J}}(t)$ moves to state $\underline{n}$ with probability $\nu_{\underline{m}\,\underline{n}}/\nu_{\underline{m}}$ as soon as it enters $\underline{m}$. To eliminate such instantaneous states inside G, we further modify $\tilde{\underline{J}}(t)$. Let $\underline{\underline{b}} = [b_{\underline{m}\,\underline{n}}]$ where $b_{\underline{m}\,\underline{n}} = \nu_{\underline{m}\,\underline{n}}/\nu_{\underline{m}}$. The good set is decomposed into subsets H and L such that

$$H = \{\underline{m} : \gamma_{\underline{m}} > 0, \quad \underline{m} \in G\}, \quad L = G \setminus H. \tag{5.6}$$

We also partition the matrix $\underline{\underline{b}}$ into submatrices

$$\underline{\underline{b}} = \begin{pmatrix} \underline{\underline{b}}_{HH} & \underline{\underline{b}}_{HL} & \underline{\underline{b}}_{HB} \\ \underline{\underline{b}}_{LH} & \underline{\underline{b}}_{LL} & \underline{\underline{b}}_{LB} \\ \underline{\underline{b}}_{BH} & \underline{\underline{b}}_{BL} & \underline{\underline{b}}_{BB} \end{pmatrix}. \tag{5.7}$$

The transition probability matrix of the corresponding replacement process $\underline{J}^{\dagger}(t)$ on $H \cup B$, eliminating the instantaneous states in G, is given by

$$\underline{\underline{b}}^{\dagger} = \begin{pmatrix} \underline{\underline{b}}^{\dagger}_{HH} & \underline{\underline{b}}^{\dagger}_{HB} \\ \underline{\underline{b}}^{\dagger}_{BH} & \underline{\underline{b}}^{\dagger}_{BB} \end{pmatrix} \tag{5.8}$$

where

$$\underline{\underline{b}}^{\dagger}_{VW} = \underline{\underline{b}}_{VW} + \underline{\underline{b}}_{VL}(\underline{\underline{I}}_{LL} - \underline{\underline{b}}_{LL})^{-1}\underline{\underline{b}}_{LW}, \text{ for } V, W \in \{H, B\}. \tag{5.9}$$

Then $\underline{J}^{\dagger}(t)$ is a Markov chain on $H \cup B$ governed by

$$\underline{\nu}^{\dagger} = [\nu^{\dagger}_{\underline{m}\,\underline{n}}]; \quad \nu^{\dagger}_{\underline{m}\,\underline{n}} = \tilde{\nu}_{\underline{m}} b^{\dagger}_{\underline{m}\,\underline{n}}, \quad \underline{m}, \underline{n} \in H \cup B. \tag{5.10}$$

We define $\underline{\nu}^{\dagger}_{D}$ as before and the matrix $\underline{\underline{a}}^{\dagger}_{\nu:HH}$ by

$$\underline{\underline{a}}^{\dagger}_{\nu:HH} = \underline{\underline{I}}_{HH} - \frac{1}{\nu}\underline{\underline{\nu}}^{\dagger}_{D:HH} + \frac{1}{\nu}\underline{\underline{\nu}}^{\dagger}_{HH}. \tag{5.11}$$

The initial state probability vector $\underline{a}$ must also be modified. We define

$$\underline{a}^{\dagger}_{H}{}^{T} = \underline{a}^{T}_{H} + \underline{a}^{T}_{L}(\underline{\underline{I}}_{LL} - \underline{\underline{b}}_{LL})^{-1}\underline{\underline{b}}_{LH}. \tag{5.12}$$

In the case of stationary or quasi-stationary measures, the replacement is similar using $\underline{e}^T$ and $(\underline{q}_G^T, \underline{0}_B^T)$ instead of $\underline{a}^T$. The results in (3.16) throuth (3.19) then provides computational procedures needed where $\underline{\underline{a}}^{\dagger}_{\nu:HH}$ and $\underline{a}^{\dagger}_{H}$ in (5.8) and (5.9) should replace $\underline{\underline{a}}_{\nu:GG}$ and $\underline{a}_G$. Namely one has

$$T^{\dagger}_{\underline{a}B} \stackrel{\mathrm{d}}{=} Z(T_{\underline{a}B}) \tag{5.13}$$

where

$$T^{\dagger}_{\underline{a}B} = \inf\{t : \underline{J}^{\dagger}(t) \in B | \underline{J}(0) = \underline{m} \text{ with probability } a_{\underline{m}}, \quad \underline{m} \in S\}, \tag{5.14}$$

and

$$P[T^{\dagger}_{\underline{a}B} > x] = \sum_{k=0}^{\infty} q_k(x) \underline{a}^{\dagger}_{H}{}^{T} \underline{\underline{a}}^{\dagger k}_{\nu:HH} \underline{1}_H. \tag{5.16}$$

§.6 A Numerical Example

In this section, we demonstrate the computational procedures described in the previous sections through a numerical example. All figures are given at the end of this section. A system we consider here consists of two processing units. Each unit fails through two phases. We define

$$S_i = \{0, 1, 2\}, \qquad 1 \leq i \leq 2, \tag{6.1}$$

where the state 2 denotes that the i-th unit is in the first phase of its up-time. The state 1 represents the operative state in the second phase and the state 0 means the failed state under repair. We assume that all relevant distributions are exponentially distributed with parameters $(\mu_i, \lambda_{i1}, \lambda_{i2})$ corresponding to S_i of (6.1).

The two processing units interact with each other in the following manner. If the i-th unit is down and other unit is in operative state k, then the whole system fails with probability p_{ik} where $i, k = 1, 2$. When the system is down while one of two units is operative, the state of the operative unit does not change until the failed unit is repaired. Let $I(t)$ be the indicator function where $I(t) = 1$ $(I(t) = 0)$ means that the systems is functioning (down) at time t. The state space $S = S_1 \times S_2 \times \{0.1\}$ then has the following 13 states:

(6.2)

linearized state number	state
0	(0,0,0)
1	(1,0,0)
2	(1,0,1)
3	(0,1,1)
4	(0,1,0)
5	(2,0,0)
6	(2,0,1)
7	(1,1,1)
8	(0,2,1)
9	(0,2,0)
10	(2,1,1)
11	(1,2,1)
12	(2,2,1)

We note that the failed state of the system due to the failure in coverage (e.g. (1,0,0), (0,1,0), (2,0,0), (0,2,0)) and failed state of the system due to failure of all modules, (0,0,0), are clearly distinguished here.

If $X_i(t)$ denotes the state of the i-th unit at time t, then the process $\underline{J}(t) = (X_1(t), X_2(t), I(t))$ is a Markov process governed by the transition rate matrix:

(6.3) $\underline{\underline{\nu}} =$

$$\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 & \mu_1 & 0 & \mu_2 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mu_2 & 0 \\
\lambda_{22} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mu_2 & 0 \\
\lambda_{12} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mu_1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mu_1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mu_2 \\
0 & 0 & \lambda_{11} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mu_2 \\
0 & p_{21}\lambda_{21} & (1-p_{21})\lambda_{21} & (1-p_{11})\lambda_{11} & p_{11}\lambda_{11} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \lambda_{21} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mu_1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mu_1 \\
0 & 0 & 0 & 0 & 0 & p_{22}\lambda_{22} & (1-p_{22})\lambda_{22} & \lambda_{11} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_{21} & (1-p_{12})\lambda_{12} & \lambda_{12} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_1 & \lambda_{11} & 0
\end{pmatrix}$$

We assume that the total job class C is decomposed into three disjoint subclasses C_j, $j = 1,2,3$. The processing capacities of the whole system and each class depending on the state of $\underline{J}(t)$ are assumed to be given below.

state number m	0	1	2	3	4	5	6	7	8	9	10	11	12
$\phi(m)$	0	0	1	2	0	0	3	3	4	0	7	8	20
$\phi_1(m)$	0	0	0.5	1	0	0	1	1	2	0	4	3	10
$\phi_2(m)$	0	0	0	1	0	0	0	1	2	0	1	3	5
$\phi_3(m)$	0	0	0.5	0	0	0	2	1	0	0	2	2	5

The values of the parameters employed in the numerical example are summarized below.

(6.4) $$(\mu_1, \lambda_{11}, \lambda_{12}) = (1, 0.4, 0.6)$$

$$(\mu_2, \lambda_{21}, \lambda_{22}) = (0.7, 0.2, 0.8)$$

(6.5) $$\underline{\underline{p}} = [p_{ik}] = \begin{pmatrix} 0.2 & 0.3 \\ 0.5 & 0.7 \end{pmatrix}.$$

It is assumed that the system starts fresh at time t so that $\underline{J}(0) = (2,2,1)$. Correspondingly the initial state probability vector is given by $\underline{a}^T = (0,0,\ldots,0,1)$.

Figure 6.1(a)

time dependent state probabilities $p_k(t)$

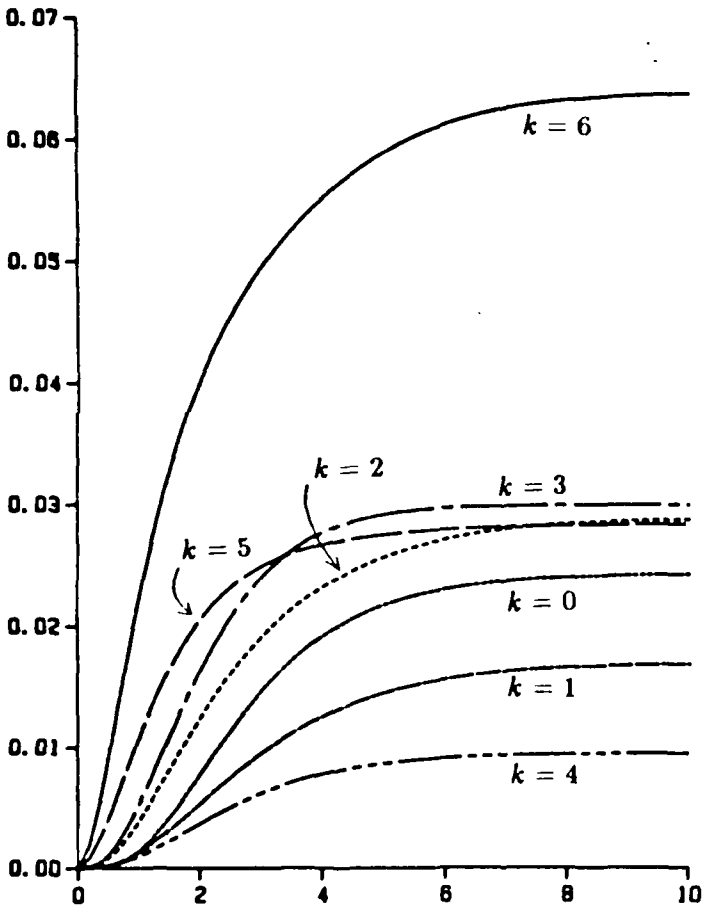


Figure 6.1(b)

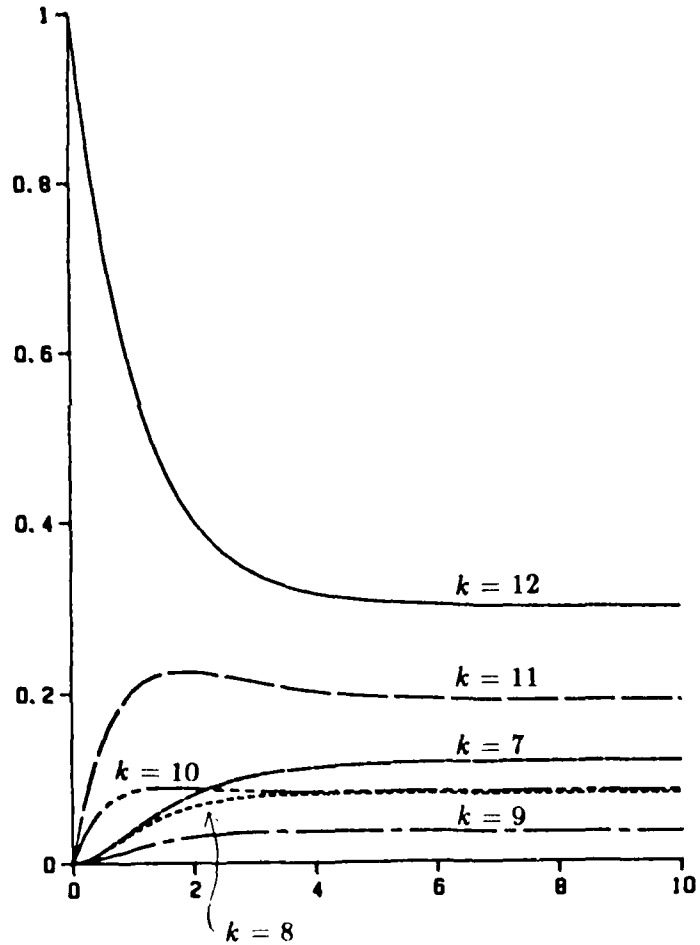


Figure 6.2(a)

the point availability $A(t|\underline{a})$

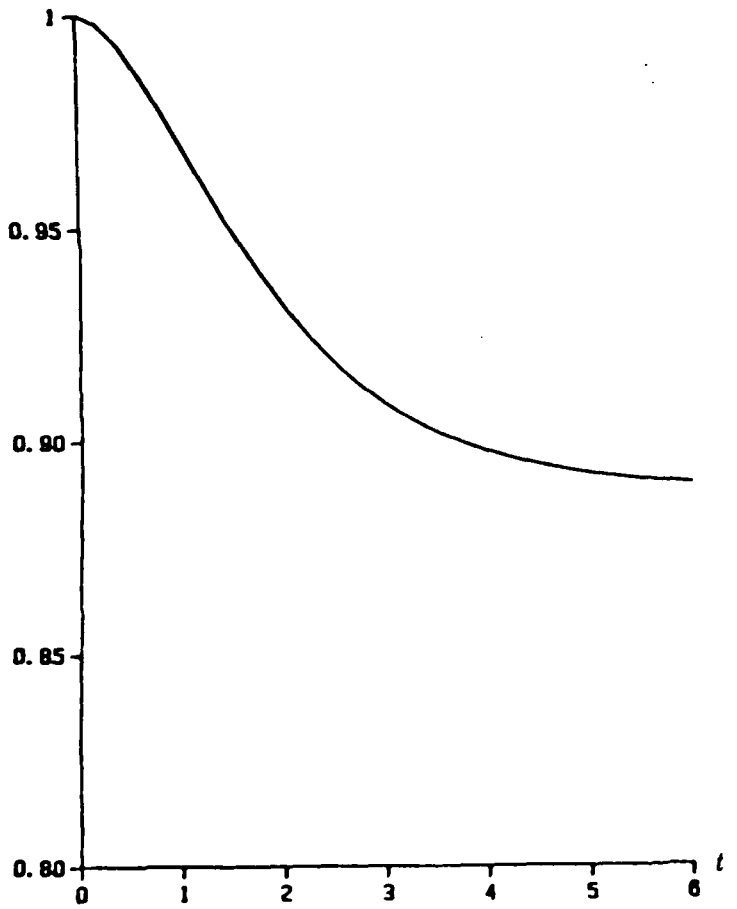


Figure 6.2(b)

the interval availability $AI(t,\tau|\underline{a})$ $t = 1.5$

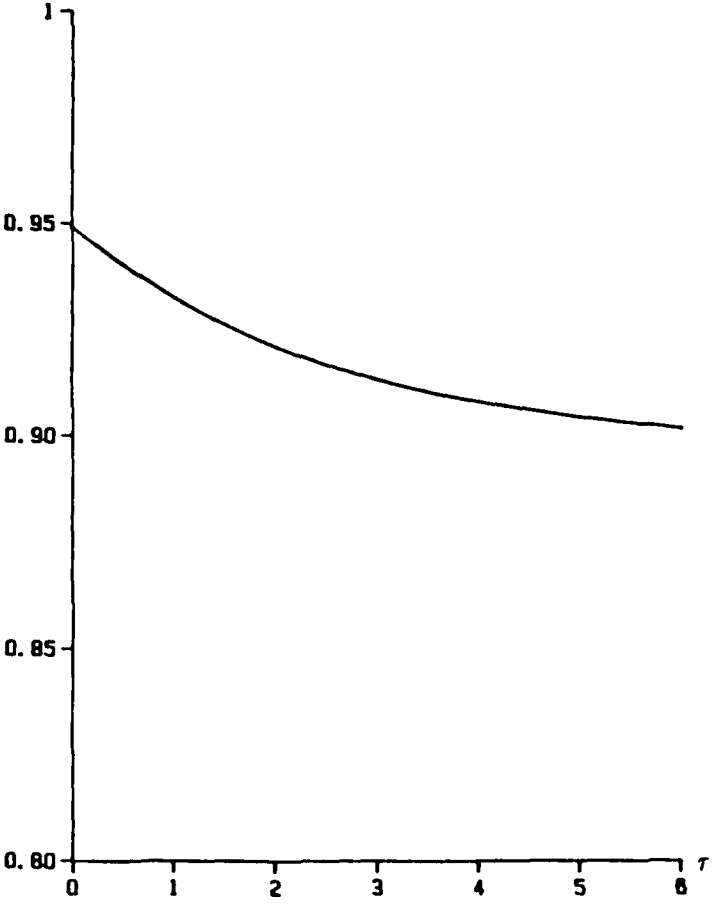


Figure 6.3(a)

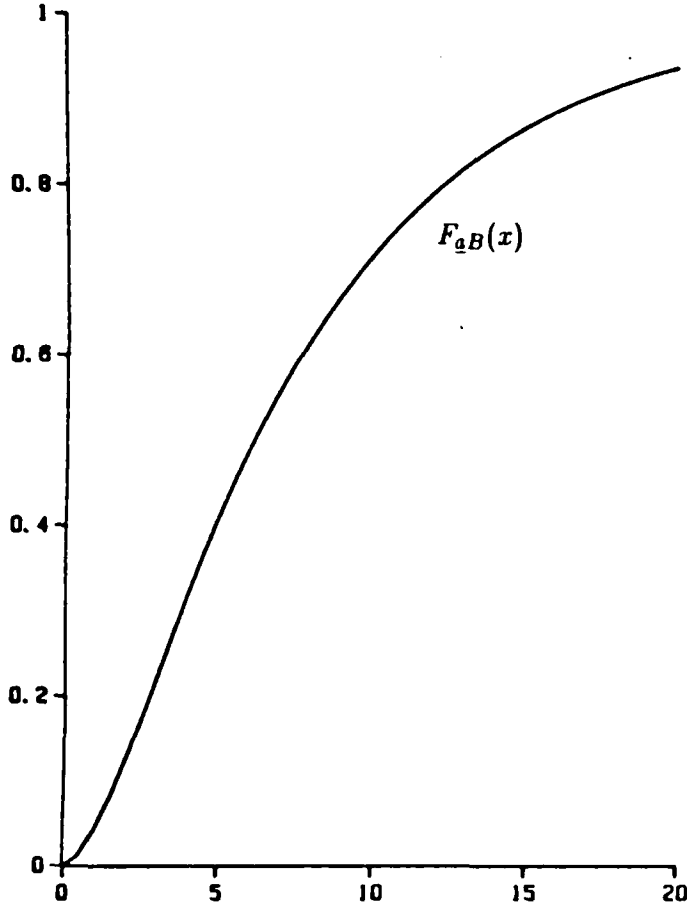


Figure 6.3(b)

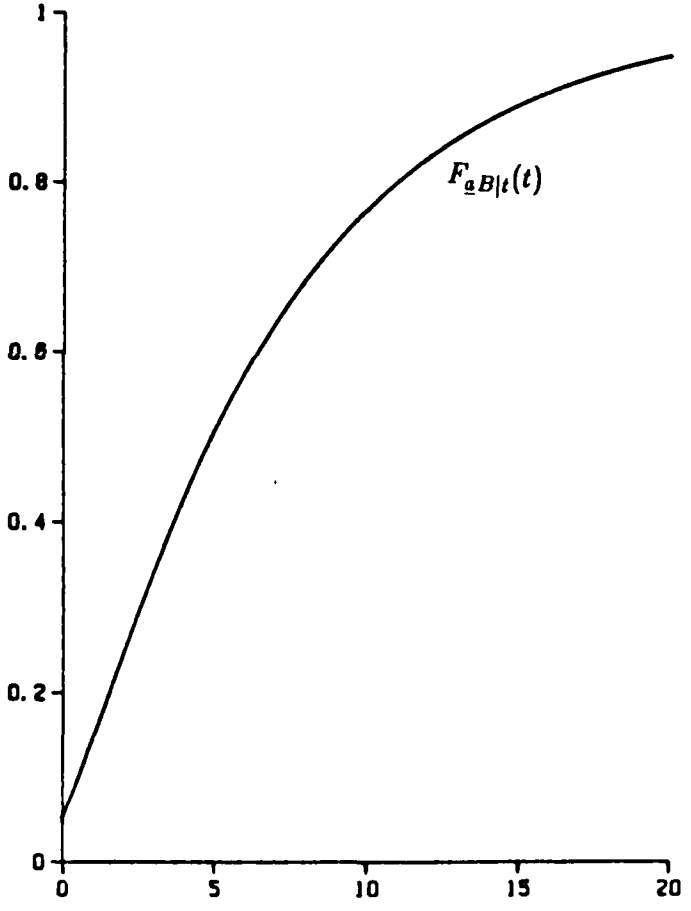


Figure 6.4

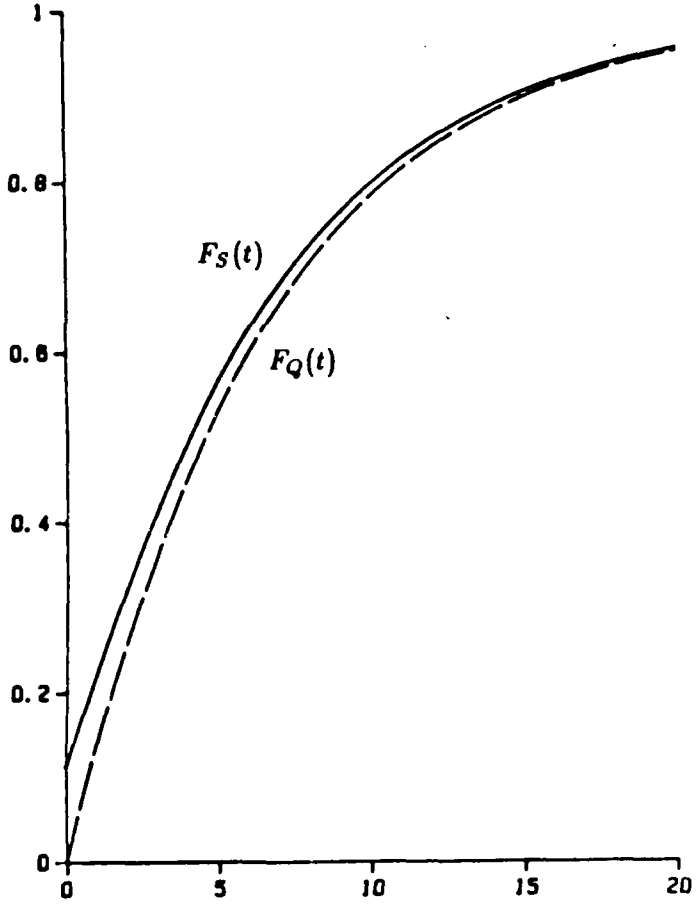


Figure 6.5

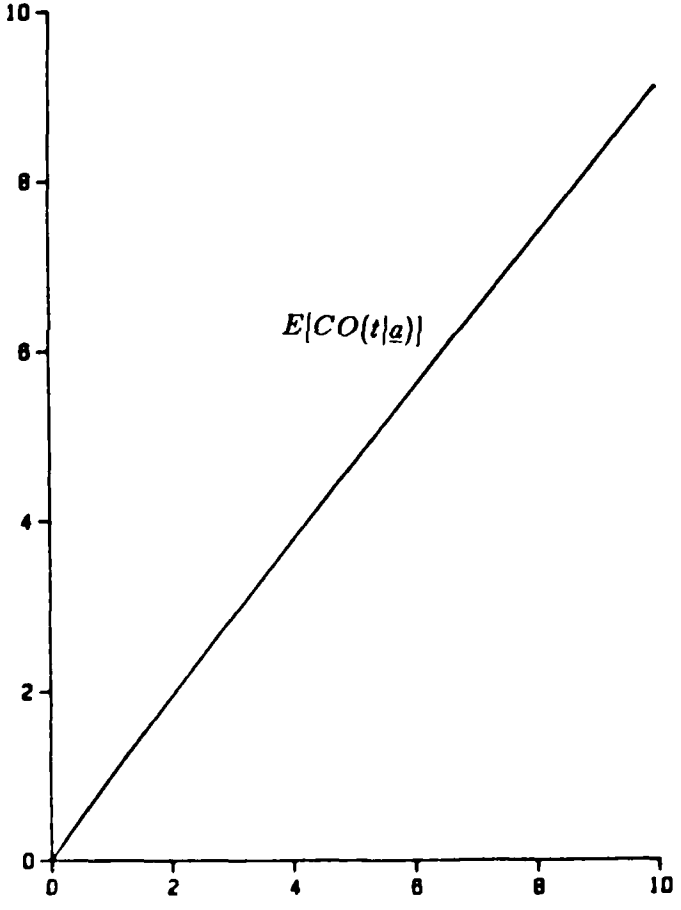


Figure 6.6

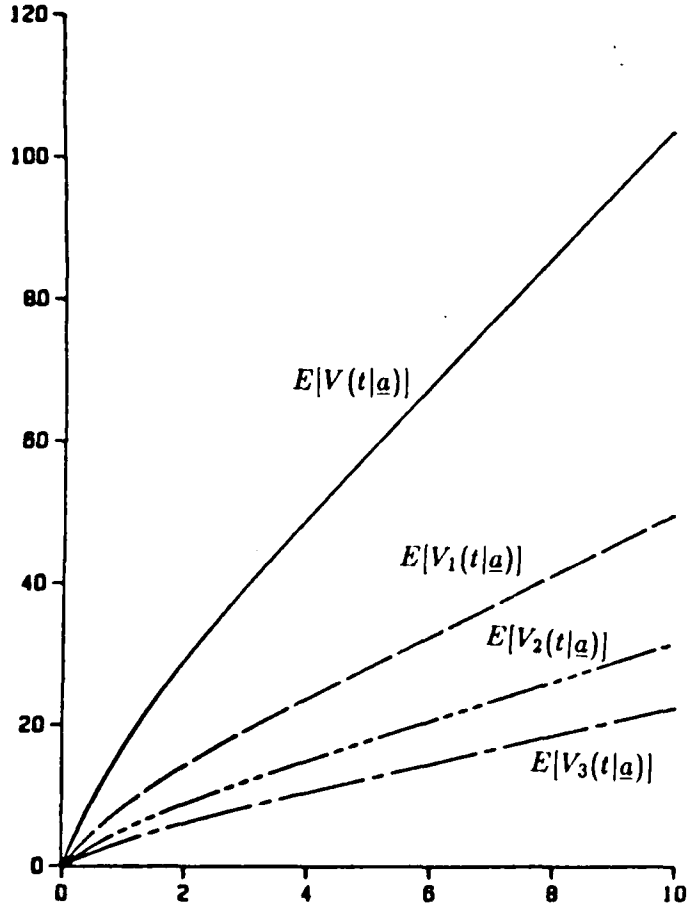


Figure 6.7(a)

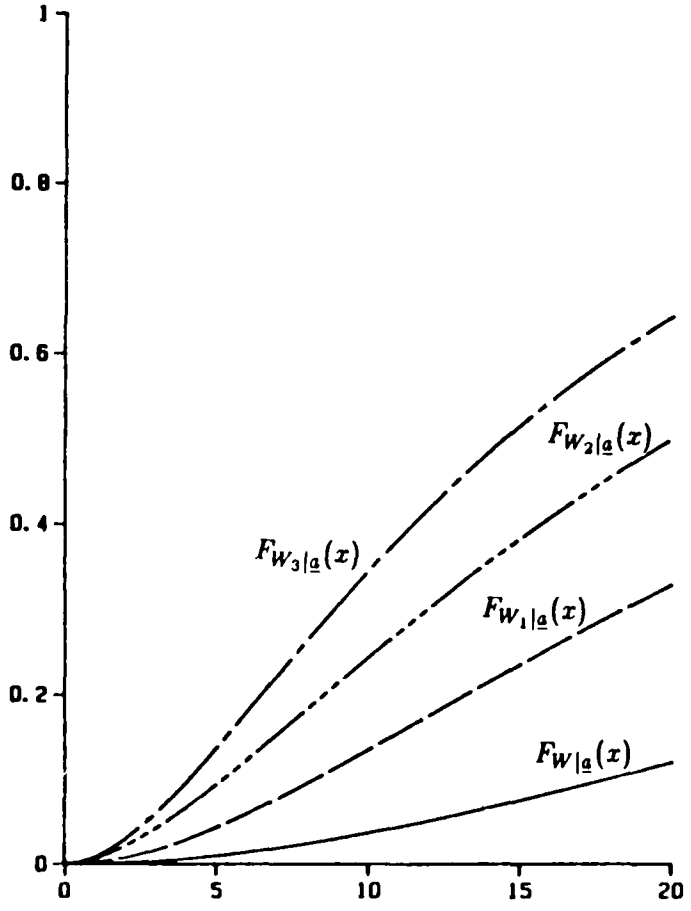


Figure 6.7(b)

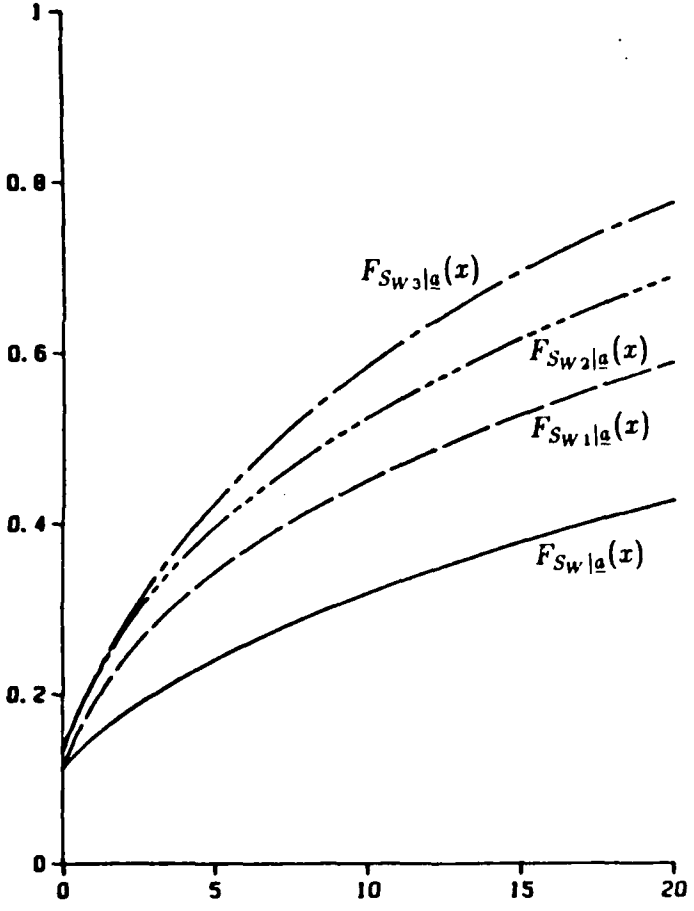
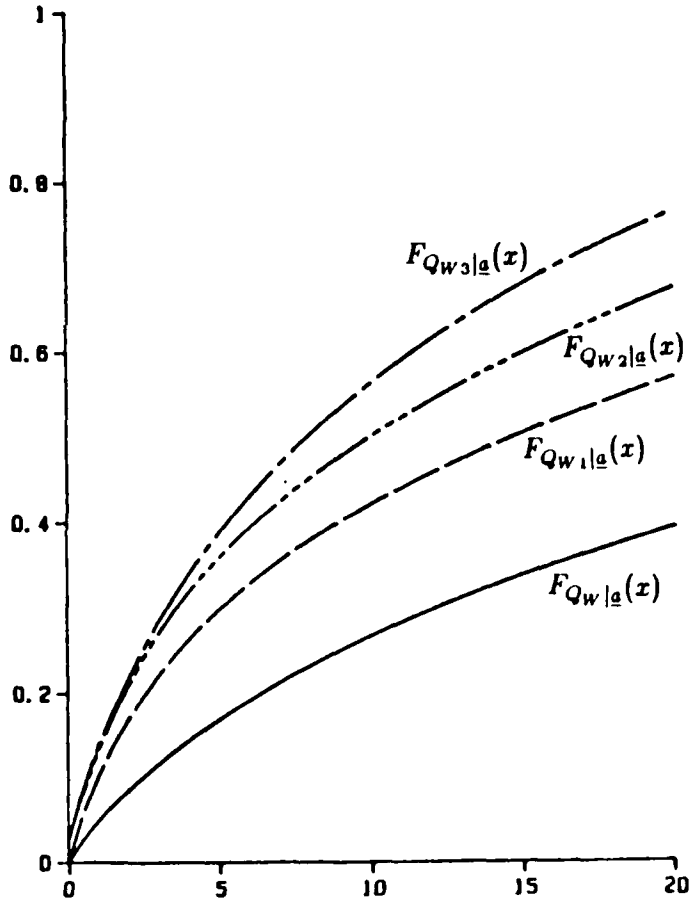


Figure 6.7(c)



Figures 6.1(a) and 6.1(b) depicts the time dependent state probabilities $p_k(t)$ for $0 \leq k \leq 6$ and $7 \leq k \leq 12$ respectively. We observe that the perfect state probability $p_{12}(t)$ decreases monotonically to its ergodic probability $e_{12} = 0.29716$, while all other state probabilities rise from zero to corresponding ergodic probabilities as $t \rightarrow \infty$, some of them monotone and some of them unimodal. The ergodicity sets in for $t \geq 4$. In Figures 6.2(a) and (b), the point availability $A(t|\underline{a})$ and the interval availability $AI(t, \tau|\underline{a})$ with $t = 1.5$ are plotted. Both decreases monotonically to common ergodic value $A_\infty = AI_\infty = 0.88757$ as $t \rightarrow \infty$.

Figure 6.3(a) and (b) illustrate the cumulative distributions $F_{\underline{a}B}(x)$ and $F_{\underline{a}B|t}(t)$ of the first passage times $T_{\underline{a}B}$ and $T_{\underline{a}B|t}$ with $t = 1.5$ respectively. We note that $T_{\underline{a}B|t}$ has a mass of $F_{\underline{a}B|t}(0+) = 1 - A(1.5)$ at the origin. Figure 6.4 depicts the cumulative distribution functions $F_S(t)$ and $F_Q(t)$ for the stationary and quasi-stationary random variables S and Q respectively. We observe that S has mass of $F_S(0+) = 1 - A_\infty$ at the origin. The values of $F_S(t)$ and $F_Q(t)$ become close for $t \geq 10$. We note that these distribution curves enable one to derive corresponding α-mission times easily. For example, τ_α^S with $\alpha = 0.8$ is 1.56. In Figure 6.5, the mean cumulative operational time of the whole system $E[CO(t|\underline{a})]$ is plotted. We see that $E[CO(t|\underline{a})]$ is almost linear having the slope $A_\infty = 0.88757$. In Figure 6.6, the mean computational capacities for each job class and entire job class are given. It is observed that all curves are concave-shaped and become quite linear for $t > 3$. Figure 6.7(a) depicts the distributions of computational capacities before the first system failure for each job class and the entire job class. Figure 6.7(b) and (c) provide corresponding curves at stationarity and quasi-stationarity. These distribution curves again enable one to evaluate the corresponding α-mission times.

REFERENCES

Arnold, T.F. (1973), "The Concept of Coverage and its Effect on the Reliability Model of a Repairable Model," *IEEE Trans. Comp.*, C-22, pp251-254.

Avizienis, A. (1978), "Fault-Tolerance: The Survival Attribute of Digital Systems," *Proc. IEEE*, vol.66, pp1109-1125.

Beaudry, M.D. (1978), "Performance-Related Reliability Measures for Computing Systems," *IEEE trans. comput.*, C-27, pp540-547.

Borgerson, B.R. and Freitas, R.F. (1975), "A Reliability Model for Gracefully Degrading and Standby-Spairing Systems," *IEEE Trans. Comp.*, C-24, pp517-525.

Bouricius, W.G. Carter, W.C. Jessep, D.C. and Schneider, P.R. (1969), "Reliability Modeling Techiniques for Self-Repairable Computer Systems," in *Proc. 24th ACM Nat. Conf.*, pp295-305.

Baxter, L.A. (1982), "Compound Availability Measures," *Nav. Res. Log. Quart*, vol.29, pp403-410.

Castillo, X. and Siewiorek, D.P. (1981), "Workload, Performance and Reliability of Digital Computer Systems," in *Proc. 1981 Int. Symp. Fault-Totelant Comput.*, Porland, ME, pp84-89.

Çinlar, E. (1969), "Markov Renewal Theory," *Adv. Appl. Prob.*, vol.1, pp123-187.

Costes, A., Landrault, C., Laprie, C. (1978), "Reliability and Availability Models for Maintained Systems Featuring Hardware Failure and Design Faults," *IEEE Trans. Comp.*, C-27, pp548-560.

Gay, F.A. and Ketelsen, M.L. (1979), "Performance Evaluation for Gracefully Degrading Systems," *Proc. 9th Ann. Int. Symp. on FT Comput.*, pp51-58.

Huslende, R. (1981), "A Combined Evaluation of Performance and Reliability for Degrading Systems," in *Proc. ACM/SIGMETRICS Conf. on Meas. and Modeling of Comput. Syst.*, Las Vegas pp157-164.

Iyer, B.R., Donatiello, L. and Heidelberger, P. (1984), "Analysis of Performability for Stochastic Models of Fault-Tolerant Systems," *IBM Research*, RC10719 (#48047).

Jewell, W.S. (1963), "Markov Renewal Programming I: Formulation, Finite Return Models," *OR*, 11, pp938-948.

Keison, J. (1974), "Markov Chain Models—Rarity and Exponentiality," *Operations Reseach Center*, ORC 74-32, College of Engineering, Univ. California, Berkeley.

Keilson, J. (1979), *Markov Chain Models—Rarity and Exponentiality*, Springer-Verlag.

Keilson, J. and Rao, S.S.(1970), "A Process with Chain Dependent Growth Rate, Part I," *J. Appl. Prob.*, vol. 7, pp699-711.

Keilson, J. and Rao, S.S. (1971), "A Process with Chain Dependent Growth Rate: The Ruin and Ergodic Problems," *Adv. Appl. Prob.*, vol. 3, pp315-338.

Koren, I. and Su, S.H.Y. (1979), "Reliability Analysis of N-Modular Redundant Systems with intermittent and Permanent Faults," *IEEE Trans. Comput.*, C-28, pp514-520.

Krishina, C.M. and Shin, K.G. (1983), "Performance Measures for Multi-Processor Controllers," *Performance '83*, A.K. Agrawala and S.K. Tripathi Eds., North-Holland, pp229-250.

Losq, J. (1976), "A Highly Efficient Redundancy Scheme: Self-Purging Redundancy," *IEEE Trans. Comput.*, C-25, pp569-578.

Makam, S.V. and Avizienis, A. (1981), "Modeling and Analysis of Periodically Renewed Closed Fault-Tolerant Systems," *Proc. 11th Int. Symp. FT comput.*, pp134-141.

Makam, S.V. and Avizienis, A. (1982), "ARIES81: A Reliability and Life-cycle Evaluation tool for Fault-Tolerant Systems," *Proc. 12th Int. Symp. FT Comput.*, pp267-274.

Masuda, Y., Shanthikumar, J.G. and Sumita, U. (1984), "A General Software Availability/Reliability Model: Numerical Exploration via the Matrix Laguerre Transformation," *Working Paper Series No. QM8425*, Univ. Rochester.

Mathur, F.P. and Avizienis, A. (1970), "Reliability Analysis and Architecture of a Hybrid Redundant Digital System: General Triple Modular Redunduncy with Self-Repair," in *Proc. 1970 SJCC, AFIPS Conf. Proc.*, vol. 36, pp375-383.

Mclean, R.A. and Neuts, M.F. (1967), "The Integral of Step Function Defined on a Semi-Markov Process," *SIAM J. Appl. Math.*, 15, pp726-738.

Meyer, J.F. (1980), "On Evaluating the Performability of Degradable Systems," *IEEE Trans. Comput.*, C-29, pp720-731.

Meyer, J.F. (1982), "Closed-Form Solutions of Performability," *IEEE Trans. Comput.*, C-31, pp648-657.

(1980)

Meyer, J.F., Furchtgott, D.G. and Wu, L.T., "Performability Evaluation of the SIFT Computer," *IEEE Trans. Comput.*, C-29, pp501-509.

Neuts, M.F. (1981), *Matrix-Geometirc Solutions in Stochastic Models*, John Hopkins Press, Baltimore, M.D.

Ng, Y.N. and Avizienis, A. (1977), "ARIES-An Automated Reliability Estimation System for Redundant Digital Structure," in *Proc. 1977 Ann. Reliability and Maintenability Symp.*, pp108-113.

Ng, Y.N. and Avizienis, A. (1980), "A Unified Reliability Model for fault-Tolerant Computers," *IEEE Trans. Comput.*, C-29, pp1002-1011.

Oda, Y., Toma, Y. and Furuya, K. (1981), "Reliability and Performance Evaluation of Self-Reconfigurable Systems with Periodic Maintenance," *Proc. 11th Int. Symp. FT Comput.*, pp142-147.

Osaki, S. and Nishino, T. (1980), *Reliability Evaluation of Some Fault-Tolerant Computer Architecture*, Springer-Verlag, NY.

Pyke, R and Schauffle, R.A. (1964), "Limit Theorem for Markov Renewal Processes," *Ann. Math. Statist.*, 35, pp1746-1764.

Seth, S.C. and Lipsky, L. (1983), "A Simplified Method to Calculate Failure Times in a Fault-tolerant Systems," *IEEE Trans. Comput.*, C-32, pp754-756.

Shanthikumar, J.G. (1983), "On a Software Availability Model with Imperfect Maintenance," *Working Paper #83-011*, Systems and Industrial Eng., Univ. Arizona.

Soneriu, M. and Suk, D.S. (1980), "Markov Model for Estimating the Reliability of Duplicated and Repairable Computing Systems," *19th Ann. Techinical Symp.*, pp87-93.

Sumita, U and Shanthikumar, J.G. (1984), "A Software Reliability Model with Multiple Error Generation and Removal," *Working Paper Series No. QM8417*, Univ. Rochester.

Trivedi, K.S. (1982), *Probability and Statistics with Reliability, Queueing and Computer Science Applications*, Prentice-Hall.

Trivedi, K.S., Gault, J.W. and Clary, J.B. (1980), "A Validation Prototype of System Reliability in Life-critical Applications," *Proc. 19th Ann. Techinical Symp.*, pp11-86.

Von Neuman, J (1956), "Probabilistic Logics and Synthesis of Reliable Organisms from Unreliable Components," in *Automated Studies*, C.E. Shannon and J. McCarthy Eds. Princeton, N.J: Princeton Univ. Press. pp43-98.


2-8